SEQUENCE LISTING

<110> Ruppert, J. Michael
Engler, Jeffrey Allen

<120> Krüppel-Like Transcriptional Factor KLF4/GKLF and Uses Thereof

<130> D6236CIP2

<141> 2004-02-11

<150> US 10/194,527
<151> 2002-07-12

<160> 8

<210> 1
<211> 13
<212> DNA
<213> Artificial sequence

<220>
<223> phosphorylated *Bst* XI adaptor

<400> 1
tcagttactc agg 13

<210> 2
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> phosphorylated *Bst* XI adaptor

<400> 2
cctgagtaac tgacaca 17

<210> 3
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> PCR primer used for recovery of proviral inserts

<400> 3
cctcactcct tctctagctc 20

<210> 4
<211> 23
<212> DNA

<213> Artificial sequence
<220>
<223> PCR primer used for recovery of proviral inserts

<400> 4
aacaaattgg actaatcgat acg 23

<210> 5
<211> 2639
<212> DNA
<213> *Homo sapiens*

<220>
<223> cDNA sequence of *GKLF*

<400> 5
tcgaggcgac cgcgacagtg gtgggggacg ctgctgagtg gaagagagcg 50

```
cagcccggcc accggaccta cttactcgcc ttgctgattg tctatttttg   100
cgtttacaac ttttctaaga acttttgtat acaaaggaac tttttaaaaa   150
agacgcttcc aagttatatt taatccaaag aagaaggatc tcggccaatt   200
tggggttttg ggttttggct tcgtttcttc tcttcgttga ctttggggtt   250
caggtgcccc agctgcttcg ggctgccgag gaccttctgg gcccccacat   300
taatgaggca gccacctggc gagtctgaca tggctgtcag cgacgcgctg   350
ctcccatctt tctccacgtt cgcgtctggc ccggcgggaa gggagaagac   400
actgcgtcaa gcaggtgccc cgaataaccg ctggcgggag gagctctccc   450
acatgaagcg acttccccca gtgcttcccg gccgccccta tgacctggcg   500
gcggcgaccg tggccacaga cctggagagc ggcggagccg gtgcggcttg   550
cggcggtagc aacctggcgc ccctacctcg gagagagacc gaggagttca   600
acgatctcct ggacctggac tttattctct ccaattcgct gacccatcct   650
ccggagtcag tggccgccac cgtgtcctcg tcagcgtcag cctcctcttc   700
gtcgtcgccg tcgagcagcg gccctgccag cgcgccctcc acctgcagct   750
tcacctatcc gatccgggcc gggaacgacc cgggcgtggc gccgggcggc   800
acgggcggag gcctcctcta tggcagggag tccgctcccc ctccgacggc   850
tcccttcaac ctggcggaca tcaacgacgt gagcccctcg ggcggcttcg   900
tggccgagct cctgcggcca gaattggacc cggtgtacat tccgccgcag   950
cagccgcagc cgccaggtgg cgggctgatg ggcaagttcg tgctgaaggc  1000
gtcgctgagc gcccctggca gcgagtacgg cagcccgtcg gtcatcagcg  1050
tcagcaaagg cagccctgac ggcagccacc cggtggtggt ggcgccctac  1100
aacggcgggc cgccgcgcac gtgccccaag atcaagcagg aggcggtctc  1150
ttcgtgcacc cacttgggcg ctggaccccc tctcagcaat ggccaccggc  1200
cggctgcaca cgacttcccc ctggggcggc agctccccag caggactacc  1250
ccgaccctgg gtcttgagga agtgctgagc agcagggact gtcaccctgc  1300
cctgccgctt cctcccggct tccatcccca cccggggccc aattacccat  1350
ccttcctgcc cgatcagatg cagccgcaag tcccgccgct ccattaccaa  1400
gagctcatgc cacccggttc ctgcatgcca gaggagccca agccaaagag  1450
gggaagacga tcgtggcccc ggaaaaggac cgccacccac acttgtgatt  1500
acgcgggctg cggcaaaacc tacacaaaga gttcccatct caaggcacac  1550
ctgcgaaccc acacaggtga gaaaccttac cactgtgact gggacggctg  1600
tggatggaaa ttcgcccgct cagatgaact gaccaggcac taccgtaaac  1650
acacggggca ccgcccgttc cagtgccaaa aatgcgaccg agcattttcc  1700
aggtcggacc acctcgcctt acacatgaag aggcattttt aaatcccaga  1750
cagtggatat gacccacact gccagaagag aattcagtat tttttacttt  1800
tcacactgtc ttcccgatga gggaaggagc ccagccagaa agcactacaa  1850
tcatggtcaa gttcccaact gagtcatctt gtgagtggat aatcaggaaa  1900
aatgaggaat ccaaaagaca aaaatcaaag aacagatggg gtctgtgact  1950
ggatcttcta tcattccaat tctaaatccg acttgaatat tcctggactt  2000
acaaaatgcc aaggggtga ctggaagttg tggatatcag ggtataaatt  2050
atatccgtga gttgggggag ggaagaccag aattcccttg aattgtgtat  2100
tgatgcaata taagcataaa agatcacctt gtattctctt taccttctaa  2150
aagccattat tatgatgtta gaagaagagg aagaaattca ggtacagaaa  2200
acatgtttaa atagcctaaa tgatggtgct tggtgagtct tggttctaaa  2250
ggtaccaaac aaggaagcca aagttttcaa actgctgcat actttgacaa  2300
ggaaaatcta tatttgtctt ccgatcaaca tttatgacct aagtcaggta  2350
atatacctgg tttacttctt tagcattttt atgcagacag tctgttatgc  2400
actgtggttt cagatgtgca ataatttgta caatggttta ttcccaagta  2450
tgccttaagc agaacaaatg tgtttttcta tatagttcct tgccttaata  2500
aatatgtaat ataaatttaa gcaaacgtct attttgtata tttgtaaact  2550
acaaagtaaa atgaacattt tgtggagttt gtattttgca tactcaaggt  2600
gagaattaag ttttaaataa acctataata ttttatctg              2639
```

SEQ 3

<210> 6
<211> 470
<212> PRT
<213> *Homo sapiens*

<220>
<223> amino acid sequence of GKLF protein

<400> 6

```
Met Ala Val Ser Asp Ala Leu Leu Pro Ser Phe Ser Thr Phe Ala
 1               5                  10                  15
Ser Gly Pro Ala Gly Arg Glu Lys Thr Leu Arg Gln Ala Gly Ala
            20                  25                  30
Pro Asn Asn Arg Trp Arg Glu Glu Leu Ser His Met Lys Arg Leu
        35                  40                  45
Pro Pro Val Leu Pro Gly Arg Pro Tyr Asp Leu Ala Ala Ala Thr
    50                  55                  60
Val Ala Thr Asp Leu Glu Ser Gly Gly Ala Gly Ala Ala Cys Gly
65                  70                  75
Gly Ser Asn Leu Ala Pro Leu Pro Arg Arg Glu Thr Glu Glu Phe
                80                  85                  90
Asn Asp Leu Leu Asp Leu Asp Phe Ile Leu Ser Asn Ser Leu Thr
            95                  100                 105
His Pro Pro Glu Ser Val Ala Ala Thr Val Ser Ser Ser Ala Ser
        110                 115                 120
Ala Ser Ser Ser Ser Ser Pro Ser Ser Ser Gly Pro Ala Ser Ala
    125                 130                 135
Pro Ser Thr Cys Ser Phe Thr Tyr Pro Ile Arg Ala Gly Asn Asp
140                 145                 150
Pro Gly Val Ala Pro Gly Gly Thr Gly Gly Gly Leu Leu Tyr Gly
155                 160                 165
Arg Glu Ser Ala Pro Pro Pro Thr Ala Pro Phe Asn Leu Ala Asp
                170                 175                 180
Ile Asn Asp Val Ser Pro Ser Gly Gly Phe Val Ala Glu Leu Leu
            185                 190                 195
```

```
Arg Pro Glu Leu Asp Pro Val Tyr Ile Pro Pro Gln Gln Pro Gln
                200                 205                 210
Pro Pro Gly Gly Gly Leu Met Gly Lys Phe Val Leu Lys Ala Ser
                215                 220                 225
Leu Ser Ala Pro Gly Ser Glu Tyr Gly Ser Pro Ser Val Ile Ser
                230                 235                 240
Val Ser Lys Gly Ser Pro Asp Gly Ser His Pro Val Val Val Ala
                245                 250                 255
Pro Tyr Asn Gly Gly Pro Pro Arg Thr Cys Pro Lys Ile Lys Gln
                260                 265                 270
Glu Ala Val Ser Ser Cys Thr His Leu Gly Ala Gly Pro Pro Leu
                275                 280                 285
Ser Asn Gly His Arg Pro Ala Ala His Asp Phe Pro Leu Gly Arg
                290                 295                 300
Gln Leu Pro Ser Arg Thr Thr Pro Thr Leu Gly Leu Glu Glu Val
                305                 310                 315
Leu Ser Ser Arg Asp Cys His Pro Ala Leu Pro Leu Pro Pro Gly
                320                 325                 330
Phe His Pro His Pro Gly Pro Asn Tyr Pro Ser Phe Leu Pro Asp
                335                 340                 345
Gln Met Gln Pro Gln Val Pro Pro Leu His Tyr Gln Glu Leu Met
                350                 355                 360
Pro Pro Gly Ser Cys Met Pro Glu Glu Pro Lys Pro Lys Arg Gly
                365                 370                 375
Arg Arg Ser Trp Pro Arg Lys Arg Thr Ala Thr His Thr Cys Asp
                380                 385                 390
Tyr Ala Gly Cys Gly Lys Thr Tyr Thr Lys Ser Ser His Leu Lys
                395                 400                 405
Ala His Leu Arg Thr His Thr Gly Glu Lys Pro Tyr His Cys Asp
                410                 415                 420
Trp Asp Gly Cys Gly Trp Lys Phe Ala Arg Ser Asp Glu Leu Thr
                425                 430                 435
Arg His Tyr Arg Lys His Thr Gly His Arg Pro Phe Gln Cys Gln
                440                 445                 450
Lys Cys Asp Arg Ala Phe Ser Arg Ser Asp His Leu Ala Leu His
```

455 460 465

Met Lys Arg His Phe

470

<210> 7

<211> 6

<212> RNA

<213> Artificial sequence

<220>

<223> polyadenylation signal

<400> 7

aauaaa 6

<210> 8

<211> 7

<212> DNA

<213> Artificial sequence

<220>

<223> minimal essential binding site for GKLF protein

<400> 8

rrggygy 7